

Convexity-constrained and nonnegativity-constrained spherical factorization in diffusion-weighted imaging

Daan Christiaens^{a,d,*}, Stefan Sunaert^{b,c,d}, Paul Suetens^{a,d}, Frederik Maes^{a,d}

^a*KU Leuven, Department of Electrical Engineering, ESAT/PSI, Leuven, Belgium*

^b*KU Leuven, Department of Imaging & Pathology, Translational MRI, Leuven, Belgium*

^c*UZ Leuven, Department of Radiology, Leuven, Belgium*

^d*UZ Leuven, Medical Imaging Research Center, Leuven, Belgium*

Abstract

Diffusion-weighted imaging (DWI) facilitates probing neural tissue structure non-invasively by measuring its hindrance to water diffusion. Analysis of DWI is typically based on generative signal models for given tissue geometry and microstructural properties. In this work, we generalize multi-tissue spherical deconvolution to a blind source separation problem under convexity and nonnegativity constraints. This spherical factorization approach decomposes multi-shell DWI data, represented in the basis of spherical harmonics, into tissue-specific orientation distribution functions and corresponding response functions, without assuming the latter as known thus fully unsupervised. In healthy human brain data, the resulting components are associated with white matter fibres, grey matter, and cerebrospinal fluid. The factorization results are on par with state-of-the-art supervised methods, as demonstrated also in Monte-Carlo simulations evaluating accuracy and precision of the estimated response functions and orientation distribution functions of each component. In animal data and in the presence of edema, the proposed factorization is able to recover unseen tissue structure, solely relying on DWI. As such, our method broadens the applicability of spherical deconvolution techniques to exploratory analysis of tissue structure in data where priors are uncertain or hard to define.

Keywords: diffusion-weighted imaging, factorization, spherical deconvolution, multi-tissue model, multi-shell HARDI, blind source separation

1. Introduction

Diffusion-weighted imaging (DWI) is a non-invasive magnetic resonance imaging technique with the unique ability to probe tissue microstructure in vivo, by measuring its hindrance to water diffusion (Le Bihan et al., 1986). The water diffusion process is sensitive to the cellular structure of the surrounding tissue, in particular the presence of cell membranes and intracellular organelles (Beaulieu, 2002). DWI is applied in both neuroscientific research and clinical practice, for studying brain organization, detecting pathology, and measuring disease progression.

The DWI signal can be represented in many ways, including the spherical harmonics (SH) basis (Frank, 2002) and the cumulant expansion (Kiselev, 2010) of which diffusion tensor imaging (DTI) (Basser et al., 1994) is a special

case. Parameters such as fractional anisotropy (FA) introduced in the context of such signal representations, are sensitive to changes in the underlying tissue microstructure. However, their interpretation at the cellular level is less straightforward.

In an effort to provide more specific measures, a myriad of models have been introduced that relate the measured signal to neural tissue structure. These models typically decompose the diffusion signal into cellular compartments, such as intra- and extra-axonal space or free water (Panagiotaki et al., 2012), weighted by their respective volume fractions. Similarly, nonnegativity-constrained spherical deconvolution (CSD) adopts a single fibre compartment of fixed anisotropy, the fibre response function (RF), which contributes linearly and independently to the DWI signal across all fibre orientations in the voxel (Tournier et al., 2004, 2007). Deconvolution then facilitates estimating the orientation distribution function (ODF) of fibres in that voxel, a metric of apparent fibre density in white matter (Raffelt et al., 2012; Dell’Acqua et al., 2013). CSD was later extended to multi-tissue (MT-)CSD (Jeurissen et al., 2014), which incorporates partial voluming with adjacent tissues that are not adequately modelled by the fibre response function (Parker et al., 2013; Roine et al., 2014). Each tissue compartment is then characterized by a fixed response function, assumed to be known a priori.

This work generalizes MT-CSD to a blind source separation

^{*}List of Abbreviations: CNSF, convexity- and nonnegativity-constrained spherical factorization; CSD, constrained spherical deconvolution; CSF, cerebrospinal fluid; DTI, diffusion tensor imaging; DWI, diffusion-weighted imaging; FA, fractional anisotropy; GM, grey matter; HARDI, high angular resolution diffusion imaging; MT-CSD, multi-tissue CSD; NMF, nonnegative matrix factorization; ODF, orientation distribution function; PVE, partial volume effect; QP, quadratic programming; RF, response function; RMS, root-mean-square; SH, spherical harmonics; SNR, signal-to-noise ratio; T1, T1-weighted image; WM, white matter.

^{*}Corresponding author

Email address: daan.christiaens@kuleuven.be (Daan Christiaens)

ration problem, akin to nonnegative matrix factorization (NMF) (Paatero and Tapper, 1994; Lee and Seung, 1999; Wang and Zhang, 2013). NMF decomposes each input vector as a nonnegative linear combination of unknown source vectors. Similarly, our approach expands the diffusion signal in a basis of response functions, adapted to the tissue structure and to the DWI data at hand. The resulting components can be associated with different normal tissue types and certain types of pathology. As such, our method strikes a balance between signal representation and tissue modelling: it seeks a decomposition that closely represents the data, subject to minimal constraints that give structural interpretation to the component basis functions.

In addition, this method addresses a very practical problem regarding multi-tissue CSD, namely estimating response functions from the data at hand. Originally, white matter (WM) fibre response functions were fitted to the DWI data in a single-fibre mask of high FA, after reorientation of the diffusion tensor eigenvectors (Tournier et al., 2004, 2007). Alternative recursive approaches have been introduced, which segment single-fibre voxels and reorient the data based on the peaks of the fibre ODFs iteratively (Tournier et al., 2013; Tax et al., 2014), or which calibrate the kernel anisotropy in each voxel separately under sparsity constraints (Schultz and Groeschel, 2013). However, these techniques do not directly generalize to other tissue types, such as grey matter (GM) and cerebrospinal fluid (CSF). Current literature therefore relies on tissue segmentation of T_1 -weighted images (T1) to define GM and CSF kernels, which requires the T1 to be aligned to the DWI data (Jeurissen et al., 2014). As this is rarely the case in practice, direct DWI tissue segmentation methods have been introduced independently and simultaneously, based on sparsity-constrained NMF (Jeurissen et al., 2015) or convexity-constrained NMF (Christiaens et al., 2015b, Appendix A) of the isotropic mean DWI signal per shell. These methods circumvent T1 requirement and are thus applicable in any reference frame without external input, but still rely on the diffusion tensor model for reorienting the DWI data in each single-fibre voxel. Here, we account for the full anisotropy of the DWI signal by extending NMF to convolution in spherical harmonics.

In related work, Xie et al. (2011) applied NMF to single-shell diffusion tensor data. Reisert et al. (2014) have introduced a more general dictionary learning method that imposes sparsity on the tissue ODFs. In contrast to their approach, we do not impose any constraints on the ODFs except for nonnegativity. Instead, we constrain the tissue RFs to be convex combinations of the data voxels. As such, physical plausibility of the tissue responses is ensured in a purely data-driven manner.

Extending our previous conference paper (Christiaens et al., 2015a), we made improvements to the initialization, the optimization, and the convergence criterion, improving the overall performance and speed of the algorithm. The accuracy and precision of our convexity- and nonnegativity-

constrained spherical factorization (CNSF) technique are evaluated in Monte Carlo simulations at various noise levels. In addition, we include results on healthy brain data, both in vivo and ex vivo, and in the presence of pathology, and show that the decomposition can be associated to known anatomy.

2. Method

2.1. Multi-tissue spherical convolution

Multi-tissue spherical convolution (Tournier et al., 2007; Jeurissen et al., 2014) assumes linear partial volume effect (PVE) to decompose the DWI signal into n tissue components, each of which is the spherical convolution of a response function (RF) and an orientation distribution function (ODF). The response function is an axially symmetric function $H_{t,b}(\theta)$ that characterizes the signal anisotropy and attenuation across b -values for each component t . Each RF is assumed to be spatially-invariant. The ODF $F_t(\theta, \phi)$ is a nonnegative function on the sphere that determines the local directionality and density of that particular component in the voxel. As such, the diffusion signal $S_b(\mathbf{g})$ in each voxel, for gradient direction \mathbf{g} and given b -value, becomes

$$S_b(\mathbf{g}) \approx \sum_{t=1}^n (H_{t,b} * F_t)(\mathbf{g}) \quad . \quad (1)$$

All functions are commonly represented in the basis of real, symmetric spherical harmonics (SH) of maximum order ℓ_{\max} (Tournier et al., 2007; Descoteaux et al., 2009; Jeurissen et al., 2014). As such, the convolution reduces to a multiplication of the coefficients of corresponding order ℓ , i.e., $s_b(\ell, m) = \sum_t \sqrt{\frac{4\pi}{2\ell+1}} h_{t,b}(\ell) f_t(\ell, m)$ with $\ell \in \{0, 2, \dots, \ell_{\max}\}$ and $m \in [-\ell, \ell]$. The response functions are axially-symmetric, and therefore constrained to the spherical harmonics of phase $m = 0$, known as zonal spherical harmonics.

For this work, we structure the SH coefficients of the DWI signal in tensor \bar{S} , indexed by the voxel v and shell b , and rewrite (1) as

$$\bar{S} \approx \bar{H} \circledast \bar{F} \quad . \quad (2)$$

In this equation, \bar{H} contains the zonal SH coefficients of the response functions, indexed by component t and shell b . \bar{F} contains the SH coefficients of the ODFs, indexed by voxel v and component t . The operator \circledast is introduced to denote spherical convolution in the SH basis, and corresponds to the matrix product of every slice $F_{\cdot,\cdot,(\ell,m)}$ with slice $H_{\cdot,\cdot,\ell}$ of corresponding order ℓ . Note that the $\ell = 0$ coefficients of \bar{F} represent the isotropic volume fraction or density of each tissue.

2.2. Convexity- and nonnegativity-constrained spherical factorization

Considering both the response functions \bar{H} and the ODFs \bar{F} as unknown, expression (2) can be seen as a NMF or blind source separation problem, in which a data matrix is decomposed as the product of a source matrix and a nonnegative weight matrix (Paatero and Tapper, 1994; Lee and Seung, 1999; Wang and Zhang, 2013). In this case, the unknown sources are the response functions of separate components, the weights are the associated ODFs, and we aim to find

$$\begin{aligned} \bar{H}^*, \bar{F}^* = \arg \min_{(\bar{H}, \bar{F})} \|\bar{S} - \bar{H} \otimes \bar{F}\|_F^2 \\ \text{s.t. } A \mathbf{f}_{v,t,\cdot} \geq 0 \end{aligned} \quad (3)$$

The matrix A evaluates the SH basis across a dense set of directions, to impose nonnegativity of the estimated ODFs denoted by vector slices $\mathbf{f}_{v,t,\cdot}$. The vector $\mathbf{f}_{v,t,\cdot}$ thus contains the SH coefficients \bar{F} at index (v, t) for all (ℓ, m) . The only parameters in this framework are the number of components n and the maximal harmonic order ℓ_{\max} of each component.

However, the solution to (3) is not unique. As illustrated in Fig. 1, the response functions \bar{H} span a n -gonal *simplicial* cone in the high-dimensional data space, radiating outwards from the origin $\mathbf{0}$. Only voxels “within” this cone are represented exactly; data points “outside” this cone give rise to the residual under minimization in (3). As such, any combination of RFs that envelops all observed data points gives rise to a zero residual, but may not necessarily be physically meaningful. Therefore, we impose a convexity constraint (Ding et al., 2010), which ensures that all sources H_t are a convex combination of the measured signal \bar{S} after reorientation. In other words, the convexity constraint ensures that all response functions are observed in the data, typically in voxels with low PVE in both spatial and angular domains. These low-PVE voxels will serve as linear basis functions that explicitly model the RFs as a function of the measured data. With the convexity constraint, the RFs are then represented as a contracted tensor-matrix product along the dimension of voxels v :

$$\bar{H} = \bar{Z} \times_v W, \quad (4)$$

such that each coefficient $h_{t,b,\ell} = \mathbf{z}_{\cdot,b,\ell} \cdot \mathbf{w}_{t,\cdot}$ with voxel weights $W \geq 0$ and $\|\mathbf{w}_{t,\cdot}\|_1 = 1$. The auxiliary tensor \bar{Z} contains the coefficients of the best fitting zonal harmonics to the data \bar{S} , across all possible orientations of a symmetry axis. These best fitting zonal harmonics are precomputed in each voxel, by reorienting the signal such that axis (θ, ϕ) coincides with the z -axis and evaluating the residual as the energy across coefficients of phase $m \neq 0$. This residual is an antipodally symmetric function on the sphere, and

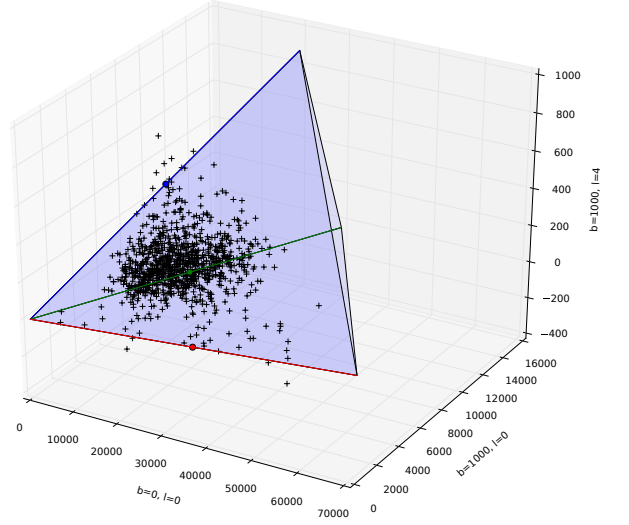


Figure 1: Illustration of the simplicial cone spanned by 3 response functions (RF) projected into a 3-dimensional subspace, shown as red, green, and blue dots. The best fitting zonal harmonic in each voxel is similarly depicted in this subspace as black crosses. Data points scattered within the simplicial cone are exactly represented as nonnegative combinations of the RFs. Data points outside this cone can not be represented exactly and give rise to a residual fitting error. The convexity constraint ensures that all RFs are convex combinations of the data points, i.e., located within the point cloud itself and typically driven towards its extremes throughout optimization.

its minimum is selected with an exhaustive search across a dense set of directions. For a corpus callosum voxel, the result typically resembles a single-fibre white matter response function. For voxels in grey matter or CSF regions, the best fitting zonal harmonic is more isotropic.

2.3. Optimization

The resulting factorization problem is computed iteratively, alternately solving for \bar{F} given \bar{H} , and for \bar{H} – implicitly represented by W – given \bar{F} . This procedure is initialized with k -means and repeated until convergence.

Initialization. The response functions \bar{H} are initialized with *spherical* k -means clustering of the best-fitting zonal harmonics \bar{Z} . Spherical k -means (Dhillon and Modha, 2001) is identical to the standard k -means algorithm (MacQueen, 1967), but uses the cosine distance instead of the Euclidian distance between data points. This cosine metric is independent of scaling effects, and instead minimizes the within-cluster angle between all datapoints. As such, spherical k -means partitions the simplicial cone of Fig. 1 in k sub-cones, making it well suited for initializing any nonnegative factorization method. Moreover, this k -means initialization obeys the convexity constraint: there exists a $W^{(0)}$ for which the initialization $\bar{H}^{(0)} = \bar{Z} \times_v W^{(0)}$.

In addition, the initialization is adapted to n response functions of given ℓ_{\max} each, by projecting all centroids to the appropriate subspace in each k -means iteration. The appropriate subspace is chosen by selecting the permutation

of centroids that minimizes the projection residual. For example, in case of $\ell_{\max} = (8, 0, 0)$ the two centroids closest to the $\ell \leq 0$ subspace are projected onto this subspace, to ensure that they represent isotropic functions. Finally, since k -means itself is randomly initialized, the entire procedure is repeated 10 times to ensure robustness, and the result of minimal residual is selected.

Solve for $\bar{F}^{(k)}$. Given response functions $\bar{H}^{(k)}$, the tissue ODFs become

$$\begin{aligned} \bar{F}^{(k)} = \arg \min_{\bar{F}} \|\bar{S} - \bar{H}^{(k)} \circledast \bar{F}\|_F^2 \\ \text{s.t. } A \mathbf{f}_{v,t,\cdot} \geq 0 \end{aligned} \quad (5)$$

When unfolding all tensors along the dimensions of shells and SH coefficients, this results in a constrained least squares problem for every voxel v . This minimization problem is solved with quadratic programming (QP) subject to non-negativity constraints on \bar{F} . Expression (5) is identical to multi-shell multi-tissue spherical deconvolution (Jeurissen et al., 2014).

Solve for $\bar{H}^{(k+1)}$. Subsequently, given ODFs $\bar{F}^{(k)}$, the new response functions become

$$\begin{aligned} W^{(k+1)} = \arg \min_W \|\bar{S} - (\bar{Z} \times_v W) \circledast \bar{F}^{(k)}\|_F^2 \\ \text{s.t. } W \geq 0 \\ \|\mathbf{w}_{t,\cdot}\|_1 = 1 \end{aligned} \quad (6)$$

This expression is cast as one global constrained least squares problem, by unfolding all tensors across voxels, shells, and SH coefficients. The optimal RF weights W are then computed with QP, using an interior point method initialized with the solution of the previous iteration.

Convergence. The alternating least squares optimization procedure is repeated until the residual $r^{(k)} = \|\bar{S} - \bar{H}^{(k)} \circledast \bar{F}^{(k)}\|_F^2$ converges to a stable minimum. The convergence criterion is met when the relative decrease in residual $(r^{(k)} - r^{(k+1)})/r^{(k)}$ is smaller than a threshold $\epsilon = 0.5\%$.

2.4. Implementation

The procedure was implemented in Python, using custom code for evaluating the SH basis and CVXOPT (Andersen et al., 2014) for QP optimization. Each shell is multiplied with the square root of its number of gradient directions, in order to equalize the fitting residual for all DWI volumes. For practical purposes, the iterative procedure is run on a subset of 1000 voxels, randomly selected across a brain mask after applying a 3-pass erosion filter. Afterwards, the ODFs are computed for the entire image based on the resulting RFs \bar{H}^* in a single run of minimization problem (5).

3. Validation

3.1. Phantom simulation

The accuracy and precision of the proposed unsupervised factorization method are evaluated and compared against supervised deconvolution with the ground-truth RFs in simulated phantom data. This phantom consists of 3 components that mimic WM, GM, and CSF, respectively represented at $\ell_{\max} = 8, 0$, and 0 . The ground-truth ODFs consist of a collection of 70 voxels containing either pure tissue (single fibre WM, GM, or CSF), 2 equally-weighted WM fibres at different crossing angles (from 0° to 90°), or WM-GM, WM-CSF and GM-CSF partial voluming (from 0 to 100%) in which WM is simulated as a 60° fibre crossing. The ground-truth RFs used in the simulations were originally estimated from selected voxels of in vivo DWI data.

Noise-free phantom DWI data are subsequently simulated with forward convolution according to (1). The DWI signal is then sampled with a uniform gradient scheme adapted to multi-shell data (Caruyer et al., 2013). This scheme contains 150 gradient directions: 5 unweighted images ($b = 0$), 20 diffusion-weighted images at $b = 1000 \text{ s/mm}^2$, 45 images at $b = 2000 \text{ s/mm}^2$, and 80 images at $b = 3000 \text{ s/mm}^2$. Finally, Rician noise is added to all data, for signal-to-noise ratio (SNR) ranging from 5 to ∞ . SNR is defined w.r.t. the mean $b = 0$ intensity in WM. At each noise level, 100 noisy data instances are generated in order to assess accuracy and precision.

3.2. Accuracy and precision

Each noisy realization of the phantom data is factorized in 3 components, one at $\ell_{\max} = 8$ and two at $\ell_{\max} = 0$. The latter two isotropic components are sorted based on their RF b -value attenuation to ensure a similar order between the estimated and ground-truth components. The mean RF of each component $\langle H_t \rangle$ is subsequently computed as the ensemble average of the estimated RFs over all noise realizations at given SNR.

Accuracy and precision of the estimated RFs are assessed with the relative root-mean-squared (RMS) difference between their coefficients H and a reference H_0

$$E_{\text{rms}}(H, H_0) = \frac{\|H - H_0\|_F}{\|H_0\|_F} \quad , \quad (7)$$

in which the Frobenius norm corresponds to the total energy over all shells according to Parseval's theorem. Accuracy is measured between the mean RF of each component $\langle H_t \rangle$ and its corresponding ground-truth RF G_t , i.e., $E_{\text{rms}}(\langle H_t \rangle, G_t)$. Precision is reported as the average error $\langle E_{\text{rms}}(H_t, \langle H_t \rangle) \rangle$ between the estimated RFs of each noise realization and their mean.

Accuracy and precision of the estimated ODFs are assessed with the error between their respective volume fractions. In addition, the accuracy and precision of the estimated ODF peaks of anisotropic component 1 are measured in the simulated WM crossing fibre region of varying

angle. To this end, the two largest local maxima of ODF 1 exceeding a threshold of 0.3 are computed with a Newton gradient-ascent method and clustered according to the reference orientations. Accuracy is then quantified as the angular bias between the average peak orientation across noise realizations and the ground truth. Precision was measured as the mean angle between each estimated fibre orientation and its respective average.

3.3. Results

The mean RF of each component at $\text{SNR} = 20$ is depicted in Fig. 2 for visual comparison to the ground truth. At this noise level, the estimated RFs are highly accurate, as evidenced by a relative RMS error $< 2\%$ and a close visual similarity in both scale and anisotropy. The bottom row of Fig. 2 shows the accuracy and precision as a function of SNR. Both accuracy and precision improve for increasing SNR and the RMS error is practically eliminated at $\text{SNR} = \infty$. At $\text{SNR} < 20$, RF accuracy reduces more strongly than precision, indicating a bias towards the Rician noise.

Secondly, Fig. 3 shows that the estimated volume fractions of each component converge towards the results of direct MT-CSD with ground-truth RFs for increasing SNR. At low SNR, CNSF provides better estimates of the true volume fractions than direct deconvolution. Hence, the reduced RF accuracy at low SNR does not deteriorate the estimated ODF volume fractions, but rather improves them thanks to the increased flexibility of adapting the RFs to the noise distribution. The residual bias in WM-GM and WM-CSF PVE voxels originates from the non-negativity constraint in both CNSF and MT-CSD, which impedes an exact representation of the SH δ -functions in the ground-truth WM ODF. The precision of all estimated volume fractions improves for increasing SNR.

Finally, the evaluation of the peak orientations of ODF 1 in Fig. 4 similarly shows that CNSF and MT-CSD are equivalent at sufficiently high SNR. For example, at $\text{SNR} = 20$ both can discriminate crossing angles $> 45^\circ$ for $\ell_{\max} = 8$. The precision of both methods is identical for all noise levels. At low SNR, direct deconvolution with ground-truth RFs has a smaller angular bias than our blind factorization approach, but is perhaps less important at this level of precision.

4. Data and results

4.1. Data and preprocessing

Dataset 1. Data of a neurologically healthy subject were provided by the WU-Minn Human Connectome Project (Van Essen et al., 2013), subject ID 100307. The diffusion data consist of 3×90 gradient directions at b -values 1000, 2000, and 3000 s/mm^2 and 18 non-diffusion-weighted images ($b = 0$), at an isotropic voxel size of 1.25 mm, and was corrected for motion, eddy current, and EPI distortions (Glasser et al., 2013). In addition, a T1 of isotropic voxel size 0.7 mm is available in the same reference frame. All

data are corrected for intensity inhomogeneity using the T1 bias field estimated with FSL FAST (Zhang et al., 2001).

Dataset 2. A multi-shell HARDI dataset of a healthy volunteer was acquired with b -values 700, 1000 and 2800 s/mm^2 along 25, 40 and 75 directions respectively, and 8 $b = 0$ images. In addition, 3 $b = 0$ images were acquired with reverse-phase encoding. The isotropic voxel size equals 2.5 mm, $\text{TR} = 7800$ ms, $\text{TE} = 90$ ms (Poot et al., 2010). The diffusion dataset was corrected for motion, eddy current, and EPI distortion using FSL EDDY and TOPUP (Andersson et al., 2003; Andersson and Sotiropoulos, 2016), as well as intensity inhomogeneity with N4 bias field estimation (Tustison et al., 2010). In addition, a T1 image is acquired at voxel size $1 \times 1 \times 1.2$ mm and rigidly coregistered to the corrected DWI.

Dataset 3. This dataset originates from a patient who suffered a grade IV glioma in the right temporal lobe, and was acquired after tumour resection. The acquisition protocol is identical to that of dataset 2, except for the absence of reverse-phase encoded $b = 0$ images. DWI images are therefore not corrected for EPI distortion and not accurately aligned to T1.

Dataset 4. DWI data of an ex vivo rhesus macaque brain were provided by the Duke Center for In Vivo Microscopy. The original acquisition, described in Calabrese et al. (2014), consisted of a high-resolution DTI dataset and a HARDI dataset of lower resolution. The former contains 12 DWI volumes at b -value 1500 s/mm^2 and a single $b = 0$ image, at an isotropic voxel size of 130 μm . The latter consists of 30 DWI volumes at $b = 4000$ s/mm^2 and one $b = 0$ image, at an isotropic voxel size of 200 μm . The high-resolution DTI dataset is subsampled to the HARDI resolution after affine registration of their corresponding $b = 0$ images.

4.2. Results

First, the presented DWI factorization method is applied to healthy human brain datasets 1 and 2. In line with the validation experiment, we select 3 components: one anisotropic component at $\ell_{\max} = 8$ and two isotropic components at $\ell_{\max} = 0$. In dataset 1, a single run in a subset of 1000 randomly selected voxels took 8 iterations until convergence, or 4 min 59 s on a standard desktop. In dataset 2, a single run took 13 iterations in 3 min 20 s. The precision of the anisotropic RF equals 3.3% in dataset 1 and 5.6% in dataset 2. Hence, this random subsampling enables fast convergence while maintaining sufficient robustness. Afterwards, deconvolution of the full image with the resulting RFs takes 15 min to a few hours, depending on the size of the data.

The resulting decomposition in RFs and ODFs is shown in Figs. 5 and 6. Figure 5 visualizes the ODFs of all components in the full images. In both datasets, anisotropic component 1 is strongly associated with WM and its ODF lobes are well aligned with the expected fibre structure.

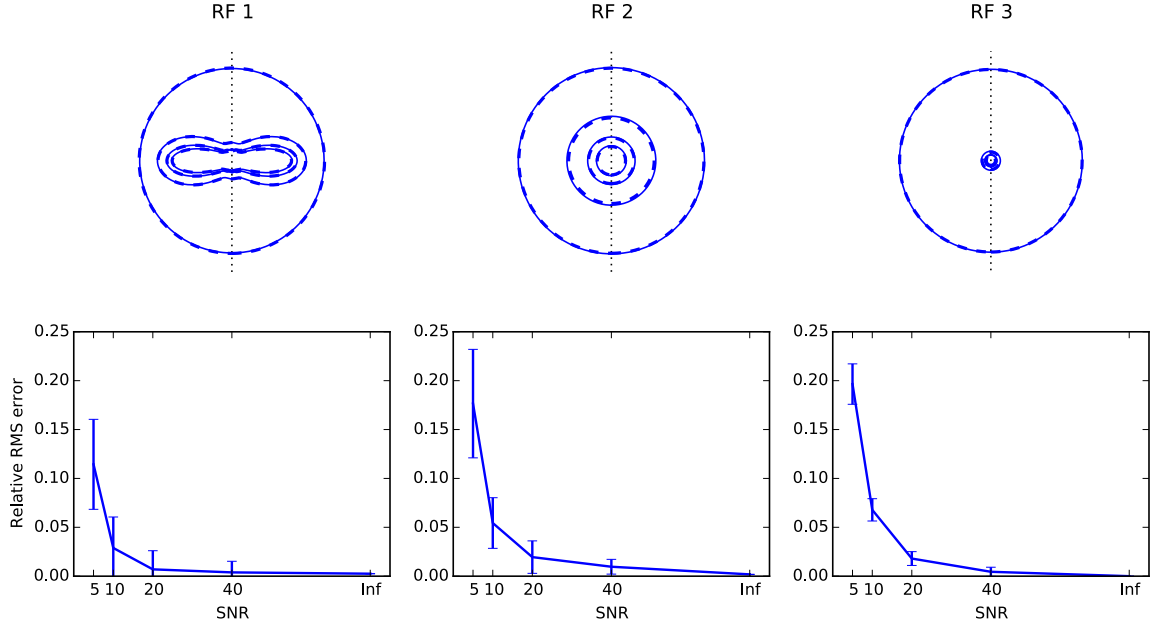


Figure 2: Response functions (RFs) in the simulated phantom data. Top row: mean RFs across all noise instances at SNR = 20 (full lines), compared to the ground truth RFs (dashed lines). RF 1 (anisotropic, $\ell_{\max} = 8$) corresponds to the simulated WM, isotropic RFs 2 and 3 correspond with the simulated GM and CSF tissues respectively. Bottom row: Accuracy \pm precision of the estimated RFs, measured with the relative RMS error to the ground-truth. Both accuracy and precision improve for increasing SNR.

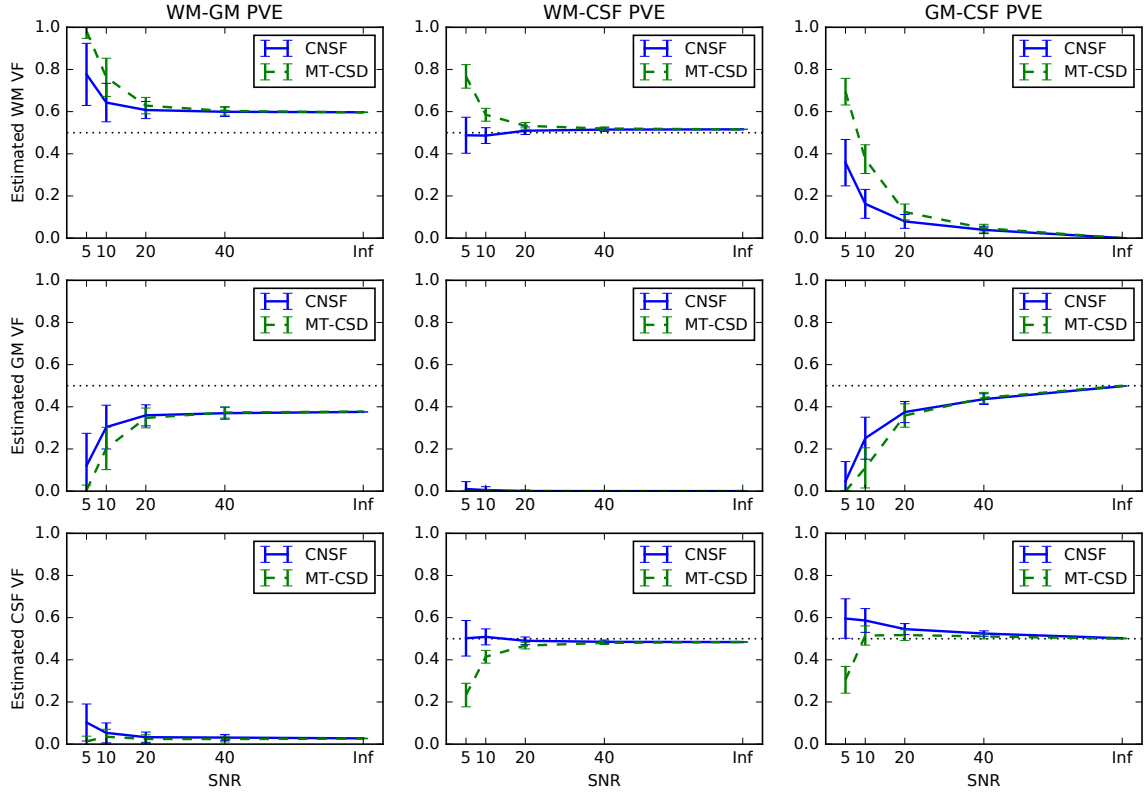


Figure 3: Accuracy \pm precision of the estimated WM, GM, and CSF volume fractions (VF) estimated with CNSF (full blue lines) and with direct MT-CSD using ground-truth RFs (dashed green lines), plotted at varying noise levels. The left column originates from a voxel with 50% WM-GM partial volume effect (PVE). The middle graphs show the estimated volume fractions in a 50% WM-CSF voxel, and the right column for a 50% GM-CSF voxel.

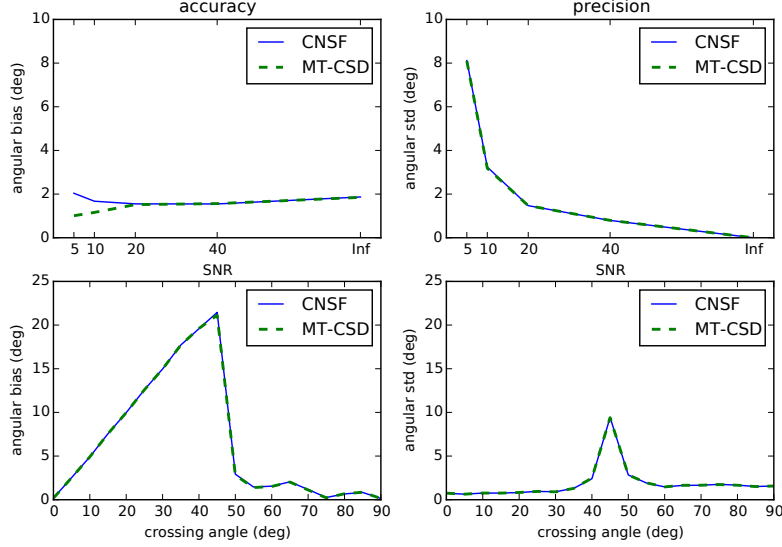


Figure 4: Accuracy (left) and precision (right) of the estimated peak orientations in ODF 1 (blue line), compared to the peak orientations of the WM fibre ODF estimated with direct deconvolution with the ground truth RFs (green dashed line). The top row plots the angular bias and precision at varying signal-to-noise ratio (SNR) in a 60° crossing. The bottom row plots these measures for different crossing angles at $\text{SNR} = 20$.

Similarly, components 2 and 3 are associated with GM and CSF contrasts. Since both components are imposed to be isotropic, their ODFs are isotropic volume fraction maps that correspond to the $\ell = 0$ SH coefficient. Note that CNSF produces these components in random order, and we manually sorted them for WM, GM, CSF correspondence. Figure 6a-b depicts the resulting RFs, which resemble the anisotropy and attenuation expected of those tissues. Figure 6c shows that the residual decreases throughout optimization and converges rapidly. Finally, Fig. 6d plots voxel weights W that represent the estimated RFs upon convergence. As shown, these weights evolve to a sparse combination of voxels, consistent with theoretical proof (Ding et al., 2010).

Next, we compare the results to MT-CSD as implemented in MRtrix3¹ (Tournier et al., 2012). A single-fibre WM RF and isotropic GM and CSF RFs are estimated from the DWI data based on a T1 tissue segmentation as described in Jeurissen et al. (2014). The WM single-fibre mask is obtained with an iterative procedure based on Tournier et al. (2013). These WM, GM, and CSF response functions are depicted in dashed lines in Fig. 6. As shown, the RFs estimated with CNSF exhibit similar attenuation across b -values, up to a scaling factor. The anisotropic RF of component 1 closely resembles the WM RF when rescaled to equalize their $b = 0$ shells. In addition, Fig. 6c shows that the residual of CNSF upon convergence is smaller than the residual of MT-CSD, indicating that a better fit of the data is obtained. In Figures 7 and 8, the ODF of component 1 is compared to the WM fibre ODF obtained with MT-CSD. Both are qualitatively very similar, showing

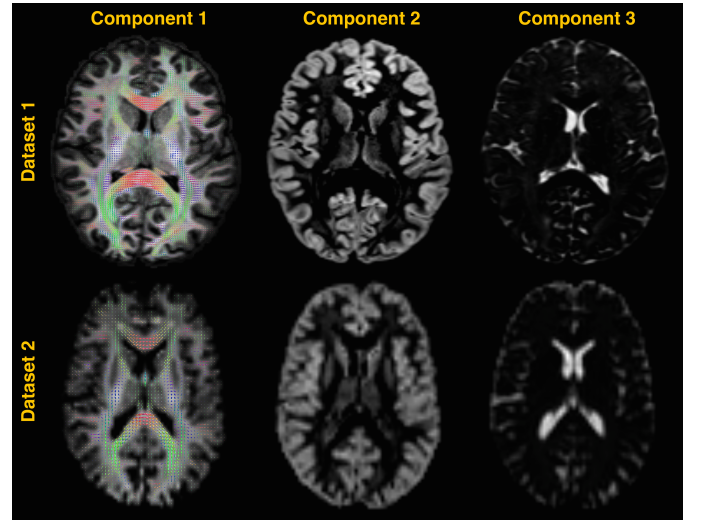


Figure 5: Factorization results with 3 components in healthy human brain datasets 1 and 2: Axial slices of the orientation distribution function (ODF) of each component. ODF 1 includes directional information associated with white matter fibre structure, ODF 2 and 3 are isotropic and are associated with GM, and CSF volume fractions.

¹J-D Tournier, Brain Research Institute, Melbourne, Australia, <https://github.com/MRtrix3/mrtrix3>

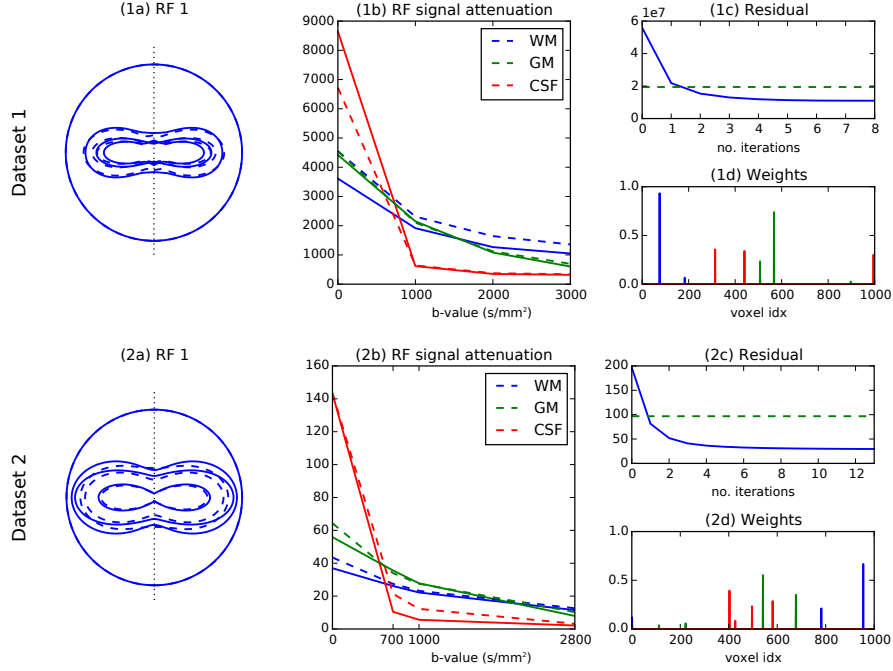


Figure 6: Factorization results with 3 components in dataset 1 (top) and in dataset 2 (bottom). (a) The anisotropic response function (RF) of component 1 (full lines) compared to the WM SF response (dashed lines) after equalizing their $b = 0$ amplitudes. (b) The RF attenuation across shells (full), compared to WM, GM, and CSF response functions (dashed). (c) The residual throughout optimization (blue curve), compared to the residual of MT-CSD (green dashed level). (d) Voxel weights encoding the estimated RFs.

fibre structure and partial voluming with adjacent tissue types. Therefore, the proposed DWI factorization method enables the benefits of multi-tissue deconvolution, without relying on T1 or external inputs.

In dataset 3, which contains residual edema surrounding the resected tumour, a decomposition in 4 components was chosen, 3 of which are constrained to isotropic RFs. As can be seen in Figs. 9 and 10, the anisotropic component is again associated to WM, whereas the first isotropic components is associated to GM and the second one to CSF. Notice how this component detects CSF in the surgical cavity, as well as in the ventricles. The third isotropic component is associated with edema in the area surrounding the resected tumour. As shown in Fig. 10, the WM fibre ODF detected in component 1 traverses this region homogeneously. While CNSF is not directly intended for lesion segmentation, this result illustrates how an unsupervised approach can discriminate pathology and adapt to outliers in abnormal data.

Finally, we demonstrate CNSF in dataset 4, which originates from an ex vivo rhesus macaque brain. Because this data contains little CSF, a factorization into two components was selected at $\ell_{\max} = 6$ and 0. As shown in Figs. 11 and 12, the resulting components are associated with WM and GM. At the exceptional spatial resolution in this dataset, this decomposition reveals WM fibres traversing distal gyri and protruding into cortical GM (Fig. 13) or branching in tree-like structure in the cerebellum (Fig. 14). These results illustrate that our method offers a practical

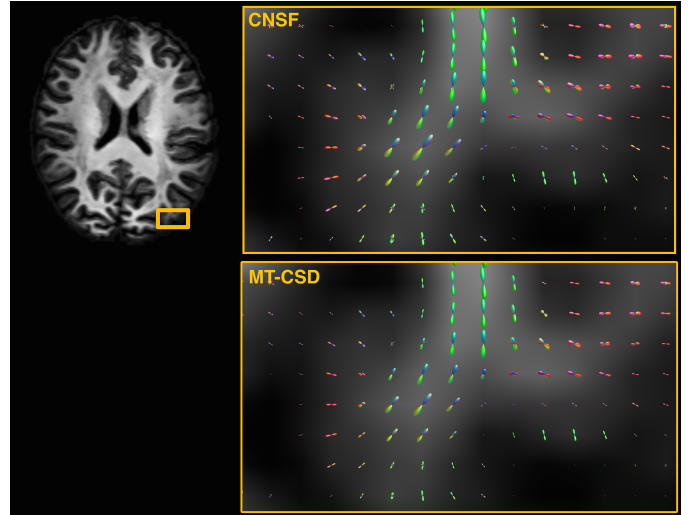


Figure 7: The ODF of the anisotropic CNSF component in dataset 1, compared to the white matter fibre ODF obtained with multi-tissue CSD. A close-up of the WM-GM interface shows fibres running through the gyrus and protruding into cortical grey matter. In both cases, explicit modelling of partial volume contamination produces a clean result with little spurious fibre directions.

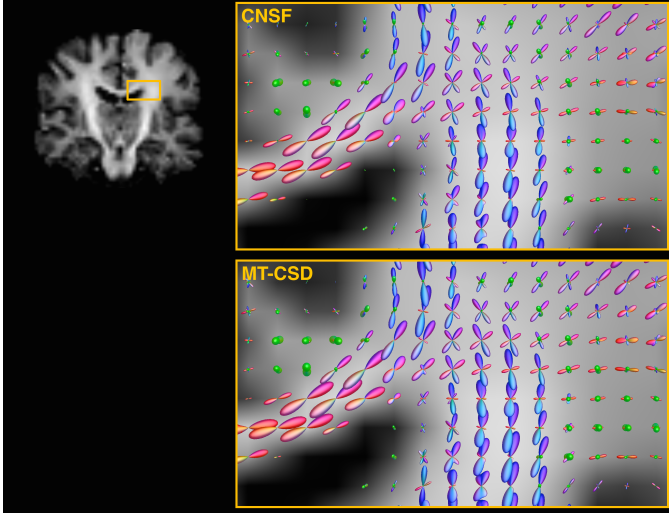


Figure 8: The ODF of the anisotropic CNSF component in dataset 2, compared to the white matter fibre ODF obtained with multi-tissue CSD. A close-up of the semioval centre shows that unsupervised CNSF factorization recovers intra-voxel fibre crossings highly similar to results of supervised MT-CSD deconvolution. In the ventricles and at the WM-CSF interface, little partial volume contamination is observed.

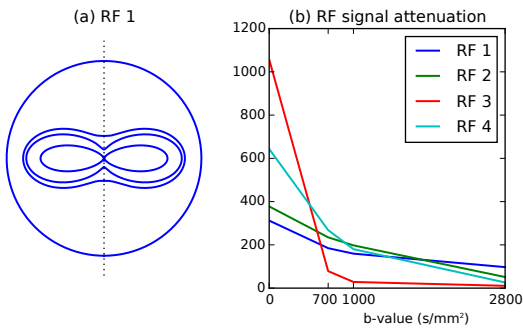


Figure 9: Response functions of 4 factorization components in dataset 3, one anisotropic component (RF 1) and three isotropic components (RF 2 – RF 4). RF 1 has the oblate shape characterizing of single-fibre white matter. RF 2 and RF 3 have signal attenuations expected of GM and CSF respectively. Finally, RF 4 has an attenuation profile between CSF and GM, associated with edema.

means of exploring tissue structure in data where no T1 or prior tissue segmentation is available.

5. Discussion

5.1. Unsupervised DWI factorization

As a direct extension of convex nonnegative matrix factorization (Ding et al., 2010) to spherical data, CNSF is an unsupervised method: it aims to discover structure in the data, without additional input. The data is represented in a generative model predicated on two minimal assumptions. First, CNSF assumes linear partial voluming between a set of tissue components, each represented by a spatially-invariant response function. Second, it assumes that these response functions are plausible, i.e., evidence of their existence must be found in the data.

MT-CSD (Jeurissen et al., 2014) also adopts the first assumption, but additionally assumes that all RFs are known a priori or estimated from the data using a prior tissue segmentation. Therefore, MT-CSD estimates tissue ODFs specifically related to the input tissue types, whereas CNSF looks for general components that best explain the data under the stated assumptions. Our results show that in many cases these components are associated with known anatomy, although this is never explicitly enforced. Both the phantom experiments and the qualitative results in real data demonstrate that CNSF factorization is on par with MT-CSD. With a fully unsupervised method and solely relying on DWI, matching the performance of its supervised counterpart is arguably the best one can aim for.

Nevertheless, due to their different interpretation CNSF and MT-CSD also serve a different purpose. CNSF is primarily suited for exploratory analysis of multi-shell DWI data in which a prior tissue segmentation is uncertain or hard to obtain. One example are cases where T1 is unavailable or not perfectly aligned to the DWI data. As demonstrated in datasets 1 and 2, CNSF successfully decomposes the DWI into WM, GM, and CSF-related contrasts, without requiring T1. A second example are cases of pathology, in which the microstructure may be altered to the extent that it is no longer accurately described by a WM-GM-CSF model. In some cases, such as our result of dataset 3, it may therefore be beneficial to include additional components. A third example are preclinical or ex vivo data or data of other organs, where the tissue structure differs from human brain. As shown in dataset 4, CNSF may discover structure in such data which is challenging to obtain with existing techniques that assume prior information.

5.2. Model selection

The main parameters to select in our approach are the number of components and the SH order ℓ_{\max} of each component. In this paper, we selected one anisotropic ($\ell_{\max} = 8$) and two isotropic ($\ell_{\max} = 0$) components for healthy human brain data, in line with Jeurissen et al. (2014). However, in other datasets it may be beneficial to

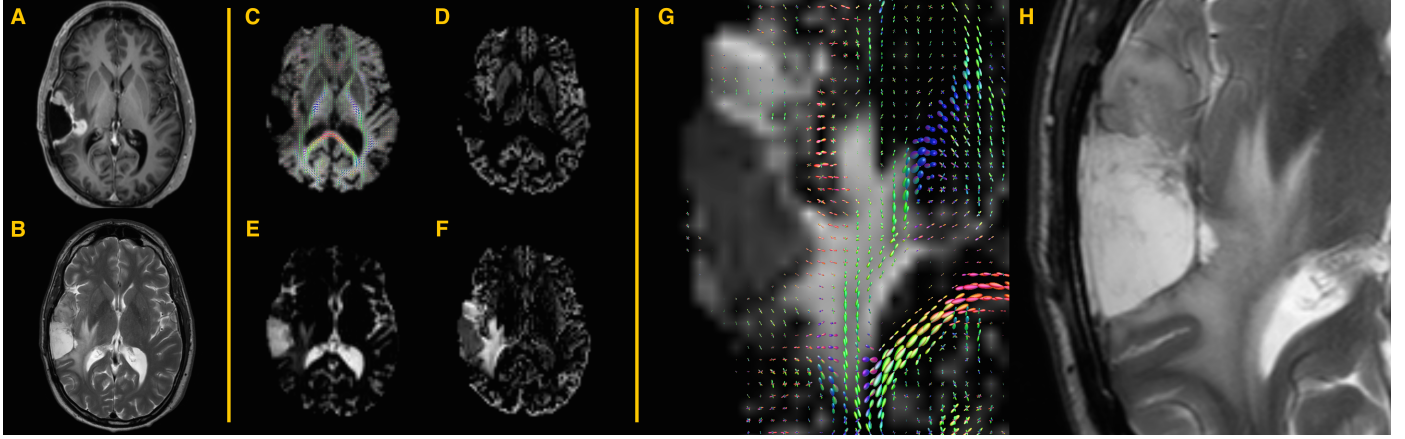


Figure 10: (A–B) T_1 - and T_2 -weighted images of dataset 3, illustrating the resected tumour and residual edema. (C–F) ODFs of components 1–4 obtained with CNSF factorization. ODF 1 recovers white matter fibre orientation. ODF 2 is associated with grey matter. ODF 3 displays CSF contrast in the ventricles and in the surgical cavity. ODF 4 highlights the edemous region surrounding the resected tumour. (G–H) A close-up of this region in ODF 1, overlaid onto component 4, shows WM fibres traversing the edemous area. A corresponding close-up of the T_2 -weighted image is provided for reference.

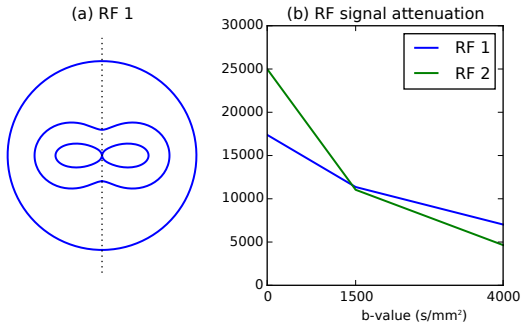


Figure 11: Response functions of 2 factorization components in dataset 4, one anisotropic component (RF 1) and one isotropic component (RF 2). RF 1 is associated with single-fibre white matter. RF 2 is associated with grey matter.

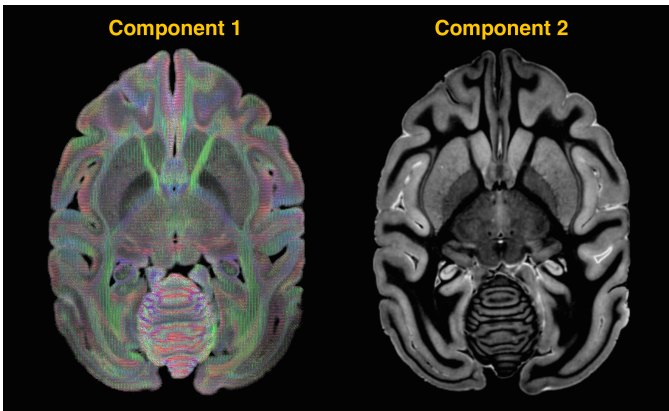


Figure 12: Factorization into 2 components in dataset 4. The ODF of anisotropic component 1 is shown on the left, and displays white matter fibre structure. The ODF of isotropic component 2, shown on the right, is primarily associated with grey matter.

use different settings. The question then arises how one should determine the optimal number of components to use. This problem is generally known as *model selection* or *rank selection*.

Model selection provides a trade-off between goodness of fit and model complexity. One approach is to use Akaike Information Criterion (AIC) (Akaike, 1974) or the Bayesian Information Criterion (BIC) (Schwarz, 1978) to select such trade-off. Another option is cross-validation (Owen and Perry, 2009). In our previous conference paper (Christiaens et al., 2015a), we applied BIC to suggest the required number of components. However, different model selection criteria are not always in agreement with each other, and which one to use remains an open question. Therefore, in this work the number of components is selected empirically, based on the nature of the data.

5.3. Future perspectives

The presented DWI factorization method lends itself to a number of applications not yet explored in the current paper. A first example is factorization of multi-modal data that includes DWI. T_1 -weighted, fluid-attenuated inversion recovery (FLAIR), MR spectroscopy metabolite contrasts, or any other scalar image can be included as additional isotropic “shells” in the input tensor \tilde{S} , provided they are co-registered with the DWI data. Such multi-modal approach may be particularly beneficial for tissue differentiation in pathology, as demonstrated in brain tumours and high-grade gliomas in particular (Sajda et al., 2004; Ortega-Martorell et al., 2012; Sauwen et al., 2015). In contrast to those earlier studies, CNSF leverages the full directional nature of the signal and assumes linearity at the level of the acquisition, rather than in derived parameters such as FA. A multi-modal approach may also “augment” single-shell DWI data to facilitate multi-tissue decomposition. Secondly, CNSF can be extended to popu-

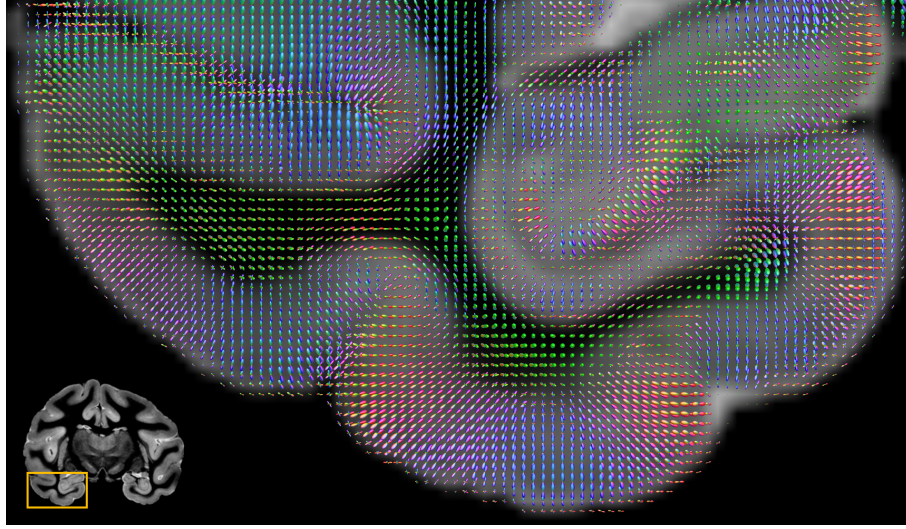


Figure 13: Coronal slice of the temporal lobe in dataset 4. The background contrast is the volume fraction of component 2. Overlaid on top is the ODF of component 1. ODF 1 shows longitudinal association fibres traversing white matter and radiating into the grey matter cortex, and recovers anisotropic tissue structure in the hippocampus.

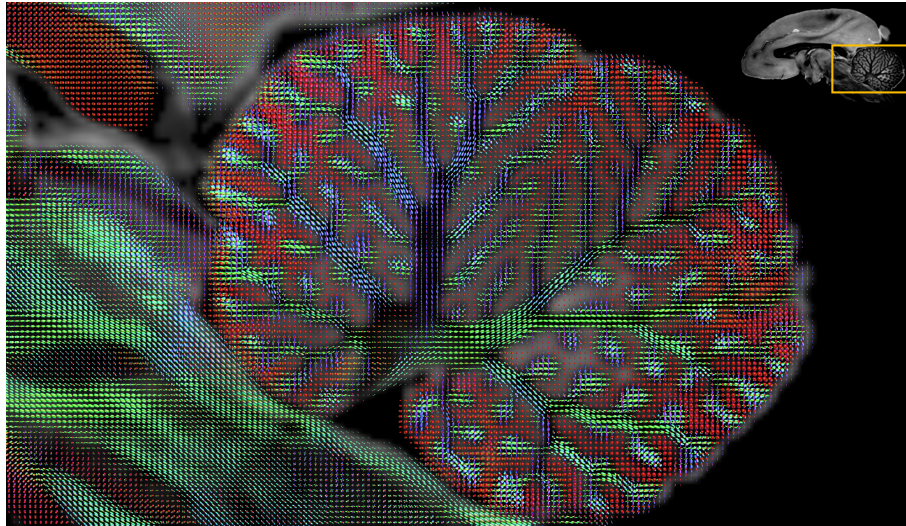


Figure 14: Sagittal slice of the cerebellum in dataset 4. The background contrast is the volume fraction of component 2. Overlaid on top is the ODF of component 1, which shows the branching structure of the arbor vitae in cerebellar white matter.

lation studies by including voxels across many subjects in the data tensor \bar{S} . As such, the resulting tissue response functions provide an optimal representation of the entire dataset, while the ODFs are quantitatively comparable across subjects. Finally, the presented DWI factorization method may have interesting applications in other organs, such as cardiac tissue or prostate tissue, in which current supervised techniques are not directly applicable.

6. Conclusion

This work introduced a generalization of multi-tissue spherical deconvolution as a blind source separation problem, formulated as convex nonnegative factorization in the SH basis. Like CSD, our approach assumes non-negativity of the tissue ODFs and spatial invariance of their RFs, but jointly optimizes the RFs instead of assuming them as known.

Acknowledgements

D. Christiaens is supported by Ph.D. grant 121013 of the Agency for Innovation by Science and Technology (IWT). This work is financially supported by KU Leuven Concerted Research Action GOA/11/006. Dataset 1 was provided the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University. Dataset 4 was provided by the Duke Center for In Vivo Microscopy, Durham, NC, USA (an NIH/NIBIB Biomedical Technology Resource Center P41 EB015897). The authors are grateful to Marco Reisert and J-Donald Tournier for interesting discussions.

References

Akaike, H., 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19, 716–723. doi:[10.1109/tac.1974.1100705](#).

Andersen, M.S., Dahl, J., Vandenberghe, L., 2014. CVXOPT: A Python package for convex optimization, version 1.1.7. URL: <http://www.cvxopt.org>.

Andersson, J.L., Skare, S., Ashburner, J., 2003. How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. *NeuroImage* 20, 870–888. doi:[10.1016/s1053-8119\(03\)00336-7](#).

Andersson, J.L., Sotiropoulos, S.N., 2016. An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *NeuroImage* 125, 1063–1078. doi:[10.1016/j.neuroimage.2015.10.019](#).

Basser, P., Mattiello, J., Le Bihan, D., 1994. MR diffusion tensor spectroscopy and imaging. *Biophysical Journal* 66, 259–267. doi:[10.1016/s0006-3495\(94\)80775-1](#).

Beaulieu, C., 2002. The basis of anisotropic water diffusion in the nervous system – a technical review. *NMR in Biomedicine* 15, 435–455. doi:[10.1002/nbm.782](#).

Calabrese, E., Badea, A., Coe, C.L., Lubach, G.R., Styner, M.A., Johnson, G.A., 2014. Investigating the tradeoffs between spatial resolution and diffusion sampling for brain mapping with diffusion tractography: Time well spent? *Human Brain Mapping* 35, 5667–5685. doi:[10.1002/hbm.22578](#).

Caruyer, E., Lenglet, C., Sapiro, G., Deriche, R., 2013. Design of multishell sampling schemes with uniform coverage in diffusion MRI. *Magnetic Resonance in Medicine* 69, 1534–1540. doi:[10.1002/mrm.24736](#).

Christiaens, D., Maes, F., Sunaert, S., Suetens, P., 2015a. Convex non-negative spherical factorization of multi-shell diffusion-weighted images, in: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (Eds.), *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Springer International Publishing, volume 9349 of *Lecture Notes in Computer Science*, pp. 166–173. doi:[10.1007/978-3-319-24553-9_21](#).

Christiaens, D., Reisert, M., Dhollander, T., Sunaert, S., Suetens, P., Maes, F., 2015b. Global tractography of multi-shell diffusion-weighted imaging data using a multi-tissue model. *NeuroImage* 123, 89–101. doi:[10.1016/j.neuroimage.2015.08.008](#).

Dell’Acqua, F., Simmons, A., Williams, S.C., Catani, M., 2013. Can spherical deconvolution provide more information than fiber orientations? Hindrance modulated orientational anisotropy, a true-tract specific index to characterize white matter diffusion. *Human Brain Mapping* 34, 2464–2483. doi:[10.1002/hbm.22080](#).

Descoteaux, M., Deriche, R., Knosche, T., Anwander, A., 2009. Deterministic and probabilistic tractography based on complex fibre orientation distributions. *IEEE Transactions on Medical Imaging* 28, 269–286. doi:[10.1109/tmi.2008.2004424](#).

Dhillon, I.S., Modha, D.S., 2001. Concept decompositions for large sparse text data using clustering. *Machine Learning* 42, 143–175. doi:[10.1023/a:1007612920971](#).

Ding, C., Li, T., Jordan, M., 2010. Convex and semi-nonnegative matrix factorizations. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32, 45–55. doi:[10.1109/tpami.2008.277](#).

Frank, L.R., 2002. Characterization of anisotropy in high angular resolution diffusion-weighted MRI. *Magnetic Resonance in Medicine* 47, 1083–1099. doi:[10.1002/mrm.10156](#).

Glasser, M.F., Sotiropoulos, S.N., Wilson, J.A., Coalson, T.S., Fischl, B., Andersson, J.L., Xu, J., Jbabdi, S., Webster, M., Polimeni, J.R., Essen, D.C.V., Jenkinson, M., 2013. The minimal preprocessing pipelines for the Human Connectome Project. *NeuroImage* 80, 105–124. doi:[10.1016/j.neuroimage.2013.04.127](#).

Jeurissen, B., Tournier, J.D., Dhollander, T., Connelly, A., Sijbers, J., 2014. Multi-tissue constrained spherical deconvolution for improved analysis of multi-shell diffusion MRI data. *NeuroImage* 103, 411–426. doi:[10.1016/j.neuroimage.2014.07.061](#).

Jeurissen, B., Tournier, J.D., Sijbers, J., 2015. Tissue-type segmentation using non-negative matrix factorization of multi-shell diffusion-weighted MRI images, in: *23rd Annual Meeting of the International Society for Magnetic Resonance in Medicine – ISMRM 2015*, p. 349.

Kiselev, V.G., 2010. The cumulant expansion: an overarching mathematical framework for understanding diffusion NMR, in: Jones, D.K. (Ed.), *Diffusion MRI*. Oxford University Press, pp. 152–168. doi:[10.1093/med/9780195369779.003.0010](#).

Le Bihan, D., Breton, E., Lallemand, D., Grenier, P., Cabanis, E., Laval-Jeantet, M., 1986. MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders. *Radiology* 161, 401–407. doi:[10.1148/radiology.161.2.3763909](#).

Lee, D.D., Seung, H.S., 1999. Learning the parts of objects by non-negative matrix factorization. *Nature* 401, 788–791. doi:[10.1038/44565](#).

MacQueen, J., 1967. Some methods for classification and analysis of multivariate observations, in: *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics*, University of California Press, Berkeley, Calif., pp. 281–297.

Ortega-Martorell, S., Lisboa, P.J.G., Vellido, A., Simões, R.V., Pumarola, M., Julià-Sapé, M., Arús, C., 2012. Convex non-negative matrix factorization for brain tumor delimitation from MRSI data. *PLoS ONE* 7, e47824. doi:[10.1371/journal.pone.0047824](#).

Owen, A.B., Perry, P.O., 2009. Bi-cross-validation of the SVD and the nonnegative matrix factorization. *Ann. Appl. Stat.* 3, 564–594. doi:[10.1214/08-aos227](#).

Paatero, P., Tapper, U., 1994. Positive matrix factorization: A

- non-negative factor model with optimal utilization of error estimates of data values. *Environmetrics* 5, 111–126. doi:[10.1002/env.3170050203](https://doi.org/10.1002/env.3170050203).
- Panagiotaki, E., Schneider, T., Siow, B., Hall, M.G., Lythgoe, M.F., Alexander, D.C., 2012. Compartment models of the diffusion MR signal in brain white matter: A taxonomy and comparison. *NeuroImage* 59, 2241–2254. doi:[10.1016/j.neuroimage.2011.09.081](https://doi.org/10.1016/j.neuroimage.2011.09.081).
- Parker, G., Marshall, D., Rosin, P., Drage, N., Richmond, S., Jones, D., 2013. A pitfall in the reconstruction of fibre ODFs using spherical deconvolution of diffusion MRI data. *NeuroImage* 65, 433–448. doi:[10.1016/j.neuroimage.2012.10.022](https://doi.org/10.1016/j.neuroimage.2012.10.022).
- Poot, D., den Dekker, A., Achten, E., Verhoye, M., Sijbers, J., 2010. Optimal experimental design for diffusion kurtosis imaging. *IEEE Transactions on Medical Imaging* 29, 819–829. doi:[10.1109/tmi.2009.2037915](https://doi.org/10.1109/tmi.2009.2037915).
- Raffelt, D., Tournier, J.D., Rose, S., Ridgway, G.R., Henderson, R., Crozier, S., Salvado, O., Connelly, A., 2012. Apparent fibre density: A novel measure for the analysis of diffusion-weighted magnetic resonance images. *NeuroImage* 59, 3976–3994. doi:[10.1016/j.neuroimage.2011.10.045](https://doi.org/10.1016/j.neuroimage.2011.10.045).
- Reisert, M., Skibbe, H., Kiselev, V.G., 2014. The diffusion dictionary in the human brain is short: Rotation invariant learning of basis functions, in: Schultz, T., Nedjati-Gilani, G., Venkataraman, A., O'Donnell, L., Panagiotaki, E. (Eds.), *Computational Diffusion MRI and Brain Connectivity*. Springer, New York. Mathematics and Visualization, pp. 47–55. doi:[10.1007/978-3-319-02475-2_5](https://doi.org/10.1007/978-3-319-02475-2_5).
- Roine, T., Jeurissen, B., Perrone, D., Aelterman, J., Leemans, A., Philips, W., Sijbers, J., 2014. Isotropic non-white matter partial volume effects in constrained spherical deconvolution. *Front. Neuroinform.* 8. doi:[10.3389/fninf.2014.00028](https://doi.org/10.3389/fninf.2014.00028).
- Sajda, P., Du, S., Brown, T., Stoyanova, R., Shungu, D., Mao, X., Parra, L., 2004. Nonnegative matrix factorization for rapid recovery of constituent spectra in magnetic resonance chemical shift imaging of the brain. *IEEE Transactions on Medical Imaging* 23, 1453–1465. doi:[10.1109/tmi.2004.834626](https://doi.org/10.1109/tmi.2004.834626).
- Sauwen, N., Sima, D.M., Cauter, S.V., Veraart, J., Leemans, A., Maes, F., Himmelreich, U., Huffel, S.V., 2015. Hierarchical non-negative matrix factorization to characterize brain tumor heterogeneity using multi-parametric MRI. *NMR in Biomedicine* 28, 1599–1624. doi:[10.1002/nbm.3413](https://doi.org/10.1002/nbm.3413).
- Schultz, T., Groeschel, S., 2013. Auto-calibrating spherical deconvolution based on ODF sparsity, in: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2013*. Springer International Publishing, pp. 663–670. doi:[10.1007/978-3-642-40811-3_83](https://doi.org/10.1007/978-3-642-40811-3_83).
- Schwarz, G., 1978. Estimating the dimension of a model. *Ann. Statist.* 6, 461–464. doi:[10.1214/aos/1176344136](https://doi.org/10.1214/aos/1176344136).
- Tax, C.M., Jeurissen, B., Vos, S.B., Viergever, M.A., Leemans, A., 2014. Recursive calibration of the fiber response function for spherical deconvolution of diffusion MRI data. *NeuroImage* 86, 67–80. doi:[10.1016/j.neuroimage.2013.07.067](https://doi.org/10.1016/j.neuroimage.2013.07.067).
- Tournier, J.D., Calamante, F., Connelly, A., 2007. Robust determination of the fibre orientation distribution in diffusion MRI: Non-negativity constrained super-resolved spherical deconvolution. *NeuroImage* 35, 1459–1472. doi:[10.1016/j.neuroimage.2007.02.016](https://doi.org/10.1016/j.neuroimage.2007.02.016).
- Tournier, J.D., Calamante, F., Connelly, A., 2012. MRtrix: Diffusion tractography in crossing fiber regions. *Int. J. Imaging Syst. Technol.* 22, 53–66. doi:[10.1002/ima.22005](https://doi.org/10.1002/ima.22005).
- Tournier, J.D., Calamante, F., Connelly, A., 2013. Determination of the appropriate b-value and number of gradient directions for high-angular-resolution diffusion-weighted imaging. *NMR in Biomedicine* 26, 1775–1786. doi:[10.1002/nbm.3017](https://doi.org/10.1002/nbm.3017).
- Tournier, J.D., Calamante, F., Gadian, D.G., Connelly, A., 2004. Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. *NeuroImage* 23, 1176–1185. doi:[10.1016/j.neuroimage.2004.07.037](https://doi.org/10.1016/j.neuroimage.2004.07.037).
- Tustison, N.J., Avants, B.B., Cook, P.A., Zheng, Y., Egan, A., Yushkevich, P.A., Gee, J.C., 2010. N4ITK: Improved N3 Bias Correction. *IEEE Transactions on Medical Imaging* 29, 1310–1320. doi:[10.1109/tmi.2010.2046908](https://doi.org/10.1109/tmi.2010.2046908).
- Van Essen, D.C., Smith, S.M., Barch, D.M., Behrens, T.E., Yacoub, E., Ugurbil, K., 2013. The WU-Minn human connectome project: An overview. *NeuroImage* 80, 62–79. doi:[10.1016/j.neuroimage.2013.05.041](https://doi.org/10.1016/j.neuroimage.2013.05.041).
- Wang, Y.X., Zhang, Y.J., 2013. Nonnegative matrix factorization: A comprehensive review. *IEEE Trans. Knowl. Data Eng.* 25, 1336–1353. doi:[10.1109/tkde.2012.51](https://doi.org/10.1109/tkde.2012.51).
- Xie, Y., Ho, J., Vemuri, B.C., 2011. Nonnegative factorization of diffusion tensor images and its applications, in: Székely, G., Hahn, H.K. (Eds.), *Information Processing in Medical Imaging: 22nd International Conference, IPMI 2011, Kloster Irsee, Germany, July 3-8, 2011. Proceedings*. Springer Berlin Heidelberg. volume 6801 of *Lecture Notes in Computer Science*, pp. 550–561. doi:[10.1007/978-3-642-22092-0_45](https://doi.org/10.1007/978-3-642-22092-0_45).
- Zhang, Y., Brady, M., Smith, S., 2001. Segmentation of brain MR images through a hidden markov random field model and the expectation-maximization algorithm. *IEEE Transactions on Medical Imaging* 20, 45–57. doi:[10.1109/42.906424](https://doi.org/10.1109/42.906424).